

Use of Screening Tests to Assess Cancer Risk and to Estimate the Risk of Adult T-Cell Leukemia/Lymphoma

by Takashi Yanagawa* and Shinkan Tokudome†

We developed methods to assess the cancer risks by screening tests. These methods estimate the size of the high risk group adjusted for the characteristics of screening tests and estimate the incidence rates of cancer among the high risk group adjusted for the characteristics of the tests. A method was also developed for selecting the cut-off point of a screening test. Finally, the methods were applied to estimate the risk of the adult T-cell leukemia/lymphoma.

Introduction

The characteristics of a screening test are usually expressed by sensitivity and specificity. The sensitivity (specificity) of a test is the probability that a person having (not having) the disease is correctly classified.

The prevalence rate of cancer is frequently estimated through the use of a screening test. The effects of the characteristics of the test in estimating disease rates have been investigated and methods are proposed for adjusting the disease rates (1,2) with the characteristics of the test assumed to be known. Methods of estimating the characteristics of the test have also been developed by several authors. Tanenbein (3) considered the use of a double sampling scheme. Hochberg (4) extended this to provide a basis for inference from general multidimensional contingency tables. Goldberg and Wittes (5) applied the capture-recapture model. Hui and Walter (6) suggested that data are available from any two populations with different prevalences and applied the method of maximum likelihood in estimating simultaneously the characteristics of the test and the prevalence rates in both populations. Yanagawa et al. (7) and Yanagawa and Kasagi (8) introduced a study design in which a person is tested repeatedly at least three different times, and they developed a method of estimating prevalence and incidence rates of disease together with

the characteristics of a test. Recently the variance of an estimator of the test characteristics was studied (9).

In this paper we consider cancers such as those caused by virus, genetic factors, and the like. The group that is infected by the virus or that inherits the genetic factors is called the high risk group. We used the situation where a screening test is employed to judge whether or not a person belongs to the high risk group. Also we determine the situation where a person is diagnosed positive (negative), if and only if the response of the test exceeds (does not exceed) the predetermined value. The value is called the cut-off point.

We first proposed a method to determine the cut-off point of the screening test. The point is determined in such a way to estimate the characteristics of the test and select the point that minimizes the estimated bias in estimating the prevalence rate. Next we developed a method of estimating the size of the high risk group in a population that adjusts for the characteristics of the screening test. The method is an extension of that proposed in Yanagawa and Tokudome (10). We next developed a method of estimating the incidence rates of cancer among the high risk group. The method adjusts for the characteristics of the screening test. There are extensive studies on this topic (11-16), but they do not discuss the characteristics of screening tests. These studies assume two binomial variates. An excellent review of these methods are given by Gart and Num (16). On the other hand, the estimation of the risk ratio is developed by Gart (17), assuming two Poisson variates. We applied the Poisson approximation in developing our estimates. Finally, the methods were applied in estimating the risk of the adult T-cell leukemia/lymphoma (ATLL) in Saga prefecture in Japan.

*Department of Mathematics, Kyushu University 33, Fukuoka 812, Japan.

†Department of Community Health Science, Saga Medical School, Saga 840-01, Japan.

Address reprint requests to T. Yanagawa, Department of Mathematics, Kyushu University 33, Fukuoka 812, Japan.

Designing a New Screening Test and Estimating Its Characteristics

We used two tests to design a new screening test and estimate its characteristics: test 1 was an expensive and sophisticated test that is used in laboratory as a standard test; test 2 was a simple and quick test, newly developed for the purpose of mass screening. A person is diagnosed positive (negative) if and only if the response of the test exceeds (does not exceed) the value of the predetermined cut-off point.

We first considered the problem of deciding the cut-off point of test 2. Usually the cut-off point is decided by testing a person by both tests, summarizing the data in a table like that in Table 1, and then selecting the point among c_1, c_2, \dots, c_{r+1} ($c_1 < c_2 < \dots < c_{r+1}$) in the table that provides test 2 with the highest association with test 1. How could this be justified? Note that, although test 1 is more reliable than test 2, it could be also a fallible test. Ideally, the cut-off point of test 2 should be decided by examining its sensitivity and specificity. The sensitivity and specificity mutually compete, and one cannot minimize them simultaneously. To circumvent this, we proposed to select the cut-off point among c_1, c_2, \dots, c_{r+1} that provided the minimum relative bias in estimating the prevalence rate.

$$D = \begin{cases} 0 & \text{if person is disease-free} \\ 1 & \text{if person has disease} \end{cases}$$

$$T_i = \begin{cases} 0 & \text{if person test negative by test } i \\ 1 & \text{if person test positive by test } i, i = 1, 2 \end{cases}$$

Denote the characteristics of test 1 by:

$$1 - \alpha = \text{pr}\{T_1 = 0 \mid D = 0\} = \text{specificity of test 1}$$

$$1 - \beta = \text{pr}\{T_1 = 1 \mid D = 1\} = \text{sensitivity of test 1.}$$

Note that α and β are called false positive and false negative rates, respectively. We introduce the corresponding quantities α_i and β_i for test 2 by:

$$\alpha_i = \text{pr}\{T_2 \geq c_i \mid D = 0\}$$

$$\beta_i = \text{pr}\{T_2 < c_i \mid D = 1\}$$

since $c_1 < c_2 < \dots < c_{r+1}$, α_i and β_i are restricted by:

$$\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_{r+1} \quad (1)$$

$$\beta_1 \leq \beta_2 \leq \dots \leq \beta_{r+1}$$

Table 1. Relationship of test 1 and test 2.

	Cut-off point	Negative	Positive	Total
Test 1	C_0	$f_{0,0}$	$f_{1,0}$	$f_{.0}$
	C_1	$f_{0,1}$	$f_{1,1}$	$f_{.1}$
Test 2	\vdots	\vdots	\vdots	\vdots
	\vdots	\vdots	\vdots	\vdots
	C_r	$f_{0,r}$	$f_{1,r}$	$f_{.r}$
	C_{r+1}	$f_{0,r+1}$	$f_{1,r+1}$	$f_{.r+1}$
	Total	$f_{0.}$	$f_{1.}$	$f_{..}$

Now when c_i is selected as the cut-off point, the probability of a person diagnosed positive by test 2 (apparent prevalence rate) is

$$P_i = \text{pr}\{T_2 \geq c_i\} \quad (2)$$

$$= \alpha_i + P_0(1 - \alpha_i - \beta_i),$$

where P_0 stands for the true prevalence rate. Therefore, the relative bias which will be introduced by test 2 with cut-off point c_i is:

$$\text{Relative bias} = |P_i - P_0|/P_0 \quad (3)$$

$$= |\alpha_i - P_0(\alpha_i + \beta_i)|/P_0.$$

Note that the usual method does not necessarily lead to the cut-off point that minimize this bias, even if test 1 is the perfect test ($\alpha = \beta = 0$), since impact of α is stronger than β in estimating the prevalence rate (2). Furthermore, note that if P_0 is small, the α and β that minimize Eq. (3) also minimize the mean square errors when the frequency is employed in estimating P_0 .

We need estimates of P_0 , α , β and α_i and β_i , $i = 1, 2, \dots, r + 1$ to assess the bias given in Equation (3). We now describe it by considering the data expressed in Table 1. The total number of the parameters involved in it is $3 + 2(r + 1) = 2r + 5$, but there is only $2r + 3$ degrees of freedom in Table 1, unless test 1 is perfect. If test 1 is perfect, all parameters are estimable. But if test 1 is not perfect, we need to introduce some reasonable restrictions on the parameters to make it estimable. We find many cases that we may supposed $\alpha_{r+1} = \beta_1 = 0$, in addition to $\alpha_1 < 1$ and $\beta_{r+1} < 1$, which perhaps reflect the effort of scientists who develop screening tests.

Put

$$f_{i.} = \sum_{j=1}^{r+1} f_{i,j}, i = 0, 1 \text{ and } f_{..} = f_{0.} + f_{1.}$$

When test 1 is perfect ($\alpha = \beta = 0$), the estimators of the parameters are given as follows:

$$p_0 = f_{.0} / f_{..}$$

$$\alpha_i = (f_{0,i} + f_{0,1} + \dots + f_{0,r+1}) / f_{..},$$

$$\beta_i = (f_{1,0} + f_{1,1} + \dots + f_{1,i-1}) / f_{..},$$

$$i = 1, 2, \dots, r + 1$$

When it is supposed that $\alpha_{r+1} = \beta_1 = 0$ and $\alpha_1 < 1$, $\beta_{r+1} < 1$, the estimators of the parameters are given by:

$$\alpha = \frac{f_{1,0}}{f_{0,0} + f_{1,0}}, \quad \beta = \frac{f_{0,r+1}}{f_{0,r+1} + f_{1,r+1}}$$

$$P_0 = \frac{1}{1 - \alpha - \beta} [1 - \alpha - (f_{0.} / f_{..})]$$

$$\alpha_1 = 1 - \frac{f_{0,0} + f_{1,0}}{(1 - P_0)f_{..}},$$

$$\beta_{r+1} = \frac{1}{P_0} \left[\frac{1}{f_{..}} \sum_{j=0}^r (f_{0,j} + f_{1,j}) - (1 - P_0) \right]$$

$$\alpha_i = 1 - \frac{1}{(1 - P_o)(1 - \alpha - \beta)f_{..}} \times \left[(1 - \beta) \sum_{j=0}^{i-1} f_{0,j} - \beta \sum_{j=0}^{i-1} f_{1,j} \right]$$

$$\beta_i = \frac{1}{P_o \beta} \left[\frac{1}{f_{..}} \sum_{j=0}^{i-1} f_{0,j} - (1 - P_o)(1 - \alpha)(1 - \alpha_i) \right],$$

$$i = 2, 3, \dots, r$$

Estimating Size of the High Risk Group in a Population by a Screening Test

Two questions frequently arise: How large is the size of the high risk group of the cancer in a population? and What is the incidence rate of the cancer in the high risk group? In this section we consider the first question; the second question will be discussed in the next section.

Let Π be the target population of size N and let Q be a subpopulation that consists of a random sample of size n from Π . Suppose that the screening test, with specificity $1 - \alpha$ and sensitivity $1 - \beta$, is undertaken to all subjects in Q to identify the high risk group, aiming to use the information for estimating the size of the high risk group in Π .

Let X be the size of the high risk group in Q , which is identified by the screening test. Suppose that X follows binomial distribution with size n and parameter P . Let Y be the size of the true high risk group in $\Pi - Q$; suppose that Y follows binomial distribution of size $N - n$ and parameter P_o . Here P_o is the true probability of a person belonging to the high risk group in Π , and P is the probability of a person diagnosed to be positive by the screening test. The latter may be called apparent rate; the term true is used to contrast it. Put

$$T = n \left\{ \frac{(X/n) - \alpha}{1 - \alpha - \beta} \right\} + Y$$

Note that T represents the size of the true high risk group in Π . We suppose that X and Y are mutually independent. The constant A that satisfies

$$pr \left\{ \frac{X - n\alpha}{n(1 - \alpha - \beta)} - A \leq \frac{T}{N} \leq \frac{X - n\alpha}{n(1 - \alpha - \beta)} + A \right\} = 1 - \alpha$$

for given α ($0 < \alpha < 1$) is obtained as

$$A = Z_\alpha \frac{N - n}{N} \left[\left\{ \frac{1}{N - n} + \frac{1}{n(1 - \alpha - \beta)^2} \right\} P_o(1 - P_o) \right]^{1/2}$$

where

$$P_o = \frac{1}{1 - \alpha - \beta} \left(\frac{X}{n} - \alpha \right) \quad (4)$$

is an estimate of the size of the high risk group adjusted for the characteristics of the test and Z_α is the upper $100(\alpha/2)\%$ point of the standard normal distribution. From this we have an approximate $100(1 - \alpha)\%$ confidence limits of the size of the high risk group in population Π adjusted for the characteristics of the screening test as follows:

$$N \left\{ \frac{X - n\alpha}{n(1 - \alpha - \beta)} \pm A \right\} \quad (5)$$

When $\alpha = \beta = 0$, the confidence limits agree with that considered in Yanagawa and Tokudome (10).

Estimating Incidence Rates of Cancer in the High Risk Group

We consider the second problem, that is, the problem of estimating the risk ratio $\mu_o = \pi / P_o$, where π is the incidence rate of the cancer in Π , and P_o is the true probability of a person belongs to the high risk group of the cancer.

Confidence Interval for μ_o

Let P again be the probability of a person diagnosed to be positive by the screening test. Let X and Z be independent binomial variates based on sample sizes n and N and parameters P and π , respectively. We consider the confidence interval of μ_o when nP and $N\pi$ are not large in a situation where the Poisson approximation to the binomial distribution is appropriate. The approximation has been employed in Gart (17) when no test characteristics are taken into account.

Using the relationship in Eq. (2), the log likelihood of x and z is obtained as:

$$l = z \ln \mu_o - N \mu_o P_o - n[\alpha + (1 - \alpha - \beta)P_o] + z \ln P_o + x \ln [\alpha + (1 - \alpha - \beta)P_o]$$

From this we have the maximum likelihood estimate of μ_o as follows:

$$\mu_o = \frac{[z / N]}{x - n\alpha / n(1 - \alpha - \beta)}$$

Note that the estimate is the unconditional maximum likelihood estimate. The conditional arguments by Gart (17) are not applicable when $\alpha > 0$ and $\beta > 0$.

Furthermore, employing the asymptotic normality of the maximum likelihood estimate, approximate confidence limits of μ_o are obtained as:

$$\mu_o \pm Z_\alpha \left[\frac{\pi(1 - \alpha - \beta)^2}{(P - \alpha)^2} \left\{ \frac{1}{N} + \frac{\pi P}{(P - \alpha)^2} \right\} \right]^{1/2} \quad (6)$$

where π is the estimated cancer incidence rate in Π , and Z_α is as defined in Eq. (4), and $P = x/n$.

If the value of P is small, say 0.05, then Eq. (6) shows that α should be extremely small, otherwise the confidence limits inflate and estimation of the risk ratio loses its validity. Note that Rogan and Gladen (1) indicate that the screening test with $\alpha = 0.05$ could be an extremely good test in practice.

Summary Risk Ratio

Suppose that there are L strata. Let X_i and Z_i be independent binomial variates based on sample sizes n_i and N_i and parameter P_i and π_i , respectively ($i=1, 2, \dots, L$). Let $\mu_{oi} = \pi_i/P_{oi}$, where P_{oi} is the true probability of a person in the stratum i belonging to the high risk group of the cancer. Putting

$$\mu_o = \mu_{o1} = \mu_{o2} = \dots = \mu_{oL}$$

we formulate the problem for obtaining a summary risk ratio which summarizes the risk ratios throughout the strata as the one of obtaining the estimate of μ_o and its confidence interval.

Now applying the same Poisson approximation as above, the log likelihood of (x_i, z_i) , $i = 1, 2, \dots, L$ is given by:

$$l = z \cdot \ln \mu_o - \mu_o \sum_{i=1}^L N_i P_{oi} - \sum_{i=1}^L n_i [\alpha + (1 - \alpha - \beta) P_{oi}] + \sum_{i=1}^L \left\{ z_i \ln P_{oi} + x_i \ln [\alpha + (1 - \alpha - \beta) P_{oi}] \right\}$$

$$\text{where } z = \sum_{i=1}^L z_i$$

From this the maximum likelihood estimate of μ_o is given by the following procedure:

Step 1: Give an initial estimate of μ_o .

Step 2: Find P_{oi} ($0 < P_{oi} < 1$) by solving the quadratic equation

$$a_i P_{oi}^2 + b_i P_{oi} + c_i = 0$$

where

$$a_i = (1 - \alpha - \beta)[\mu_o N_i + n_i(1 - \alpha - \beta)]$$

$$b_i = \alpha[\mu_o N_i + n_i(1 - \alpha - \beta)] - (x_i + z_i)(1 - \alpha - \beta)$$

$$c_i = -\alpha z_i$$

Step 3: Obtain

$$\mu_o = \frac{z}{\sum_{i=1}^L N_i P_{oi}}$$

Step 4: Repeat Step 1–Step 3 until convergence.

An approximate $100(1-\alpha)\%$ confidence limit of μ_o is given by:

$$\mu_o \pm Z_{\alpha} [\mu_o^2 / z]^{1/2}$$

Application to Estimating Risk of ATLL

The adult T-cell leukemia/lymphoma (ATLL) is the disease proposed recently (18). The disease is caused by retrovirus termed human T-lymphotropic virus type I (HTLV-I). The disease is endemic in Japan and is contracted from HTLV-I carriers. The virus may be transmitted from mother to child and between husband and wife. Further studies have shown that HTLV-I can be transmitted by blood transfusion. The authors were involved in the project of estimating the risk of ATLL in Saga prefecture in Kyushu. The high risk group of the malignant neoplasm in this study is the one with HTLV-I carriers. The cases of ATLL were obtained from the cancer registry and also from related records in the prefecture. The number of the carriers is estimated by the examination of the sera of blood donors. It is important to study the bias that might have been introduced by the use of donor's blood, but we could not quantify it for various practical reasons. The epidemiologic results of the whole study are published elsewhere (19).

Selecting the Cut-off Point

The anti-adult T-cell leukemia/lymphoma-associated antigen (ATLA) of serum has been examined in the laboratory by an indirect immunofluorescence (IF) test using acetone-fixed MT-1 cells on slides as ATLA (20). When IF dilution of serum of 1:5 or higher gives a positive immunofluorescence, the serum is judged to be IgG anti-ATLA-positive. A simple and sensitive gelatin particle agglutination (PA) procedure has been developed for mass screening of the antibody to HTLV-I in human sera. The procedure in the PA test is much simpler than that in the IF test or in enzyme-linked immunosorbent assays (ELISA) because it involves only a one-step reaction of antigen and antibody, whereas more steps are required in the other tests. The titer of the antibody in the PA test is expressed as the reciprocal of the greatest serum dilution that gives positive reactions.

Several authors (21) suggest selecting 2^4 as the cut-off point of the titer in the PA test, namely to judge the final serum dilution of $1:2^4$ or higher showing agglutination as positive. We examine this cut-off point based on the data in Table 2, which are given by Kamihira et al. (22). The data have been collected in a different sampling scheme, but the results in previous papers (21,23) indicate that we can consider the table as an example of Table 1. Table 2 lists the results of 19,816 sera tested by both tests. The two estimates of the parameters from Table 2 that were obtained by assuming the perfection of the IF test and the restriction previously described above, respectively, are shown to be identical in these data; these estimates are listed in Table 3 together with the relative bias obtained from Eq. (3). The table shows that when 2^4 is selected as the cut-off point, the bias is 0.123; when 2^6 is selected, the bias is 0.007, which is the smallest among the four.

Table 2. The IF test and the PA test.^a

	Cut-off point titer	Negative	Positive	Total
IF test	$< 2^4 \times$	18,809	0	18,809
	$2^4 \times$	37	0	37
PA test	$2^6 \times$	63	16	79
	$2^6 \times$	10	21	31
	$> 2^6 \times$	0	860	860
Total		18,919	897	19,816

^a From Kamihira et al. (22).**Table 3. Estimated characteristics and the relative bias of the PA test for each cut-off point, estimated from the data in Table 2.^a**

Cut-off point titer	α	β	Relative bias
$2^4 \times$	0.00581	0	0.123
$2^5 \times$	0.00386	0	0.081
$2^6 \times$	0.00053	0.01784	0.007
$> 2^6 \times$	0.00000	0.04125	0.041

^a α = false positive rate; β = false negative rate.**Table 4. Estimates of the size of the HTLV-I carriers and its 95% confidence intervals in Saga.**

Parameter	Age			
	20-30	30-40	40-50	50-60
Male				
Population (N)	51,486	62,229	53,506	52,736
Blood donors (n)	9800	10830	4916	2660
Estimated number of the carriers	903	1816	2255	3203
Upper limit	1026	1999	2547	3678
Lower limit	781	1634	1963	2728
Female				
Population (N)	59,100	63,706	59,144	59,946
Blood donors (n)	6656	3677	2630	2073
Estimated number of the carriers	1081	2101	2671	4062
Upper limit	1263	2465	3139	4712
Lower limit	898	1737	2204	3413

Thus we selected 2^6 as the cut-off point of the PA test in our study; this leads to the estimated false positive and negative rates of the PA test as follows:

$$\alpha = 0.00053, \beta = 0.01784.$$

Estimating Number of Carriers

Table 4 shows the data and the estimates of the number of carriers and its confidence intervals. The data are collected through the PA test with the cut-off point 2^6 , and the estimates are obtained by assuming that $\alpha = 0.00053$ and $\beta = 0.01784$ are known constants. To see the impact of the characteristics of the test further, we obtained the estimates of the size of the carriers and its confidence intervals for several values of α and β in females based on the data in Table 4. These estimates are summarized in Table 5. The table shows substantial effect of the characteristics; in particular, the impact on the younger age group is remarkable.

Table 5. Estimates of the carriers in females for selected values of α and β based on the data in Table 4.

α, β	Confidence interval ^a	Age			
		20-30	30-40	40-50	50-60
0.005, 0.06		852	1901	2522	3978
	UL	1022	2264	2998	4652
	LL	682	1539	2205	3304
0.005, 0.10		890	1986	2635	4156
	UL	1071	2372	3142	4874
	LL	710	1600	2127	3439
0.017, 0.06		95	1098	1786	3251
	UL	153	1379	2194	3872
	LL	37	817	1377	2630
0.017, 0.10		99	1148	1867	3398
	UL	161	1447	2302	4060
	LL	37	848	1431	2736

^a UL = upper level of the 95% confidence interval; LL = lower level of the 95% confidence interval.**Table 6. Incidence rates of ATLL among the HTLV-I carriers in Saga ($\times 10^{-5}$).**

Parameter	Age			Summary
	30-40	40-50	50-60	
Male				
Population (N)	62,229	53,506	52,736	
Deaths: 3 years	0	8	11	
Incidence rate	0	118.2	114.5	116.1 ^a (63.9-168.3) ^b
Female				
Population (N)	63,706	59,144	59,946	
Deaths: 3 years	3	5	7	
Incidence rate	47.6	62.4	57.4	56.6 ^c (27.9-85.2) ^b

^a Summary of age 40-60 years.^b 95% Confidence interval.^c Summary of age 30-60 years.

Estimating Incidences of ATLL among Carriers

Table 6 lists the number of cases of ATLL for 3 years and also shows the population in Saga prefecture between the ages of 30 to 60. No cases of ATLL are reported in ages 20 to 40 in males and age 20 to 30 in females. Table 6 also lists the estimates of the incidence rates of ATLL among the HTLV-I carriers in males and females and its summary rates. The estimates are obtained by the method developed in the text. It was supposed that $\alpha = 0.00053$, $\beta = 0.01784$ are known constants. To study the impact of the characteristics of the test, estimates were obtained for several values of α and β in females based on the data in Tables 4 and 6. The estimates are summarized in Table 7. The table shows substantial impact for large values of α . Since there is no incidence of ATLL for the group ages 20 to 30 on which the impact is the greatest, as was discussed previously, the table shows that the impact of the characteristics of the screening test on the summary risk ratio is not so strong as the estimation of the size of the carriers. Note that if the values of α exceed 0.032, the

Table 7. Estimated incidence rates of ATLL among the carriers in females for selected values of α and β based on the data in Table 6 ($\times 10^{-5}$).

α, β	Age			Summary
	30-40	40-50	50-60	
0.005, 0.06	52.6	66.1	58.7	59.5 (29.4-89.6) ^a
0.005, 0.10	50.3	63.3	56.1	57.0 (28.1-85.8) ^a
0.017, 0.06	91.1	93.3	71.8	81.5 (40.3-122.9) ^a
0.017, 0.10	87.1	89.3	68.7	78.0 (38.5-117.5) ^a
0.032, 0.10	150.2	190.1	96.4	141.3 (69.8-212.8) ^a

^a95% Confidence interval.

summary risk ratio in age group 30 to 60 is not obtainable because of the strong impact, in particular, on age 30 to 40.

REFERENCES

- Rogan, W. J., and Gladen, B. C. Estimating prevalence from the results of a screening test. *Am. J. Epidemiol.* 107 (1978).
- Yanagawa, T., and Gladen, B. Estimating incidence and remission from diagnostic test. *Am. J. Epidemiol.* 119: 1015-1023 (1984).
- Tenenbein, A. A double sampling scheme for estimating from binomial data with misclassification. *J. Am. Stat. Assoc.* 65: 1350-1361 (1970).
- Hochberg, Y. On the use of double sampling schemes in analyzing categorical data with misclassification errors. *J. Am. Stat. Assoc.* 72: 914-921 (1977).
- Goldberg, J. D., and Wittes, J. T. The estimation of false negatives in medical screening. *Biometrics* 34: 77-86 (1978).
- Hui, S. L., and Walter, S. D. Estimating the error rates of diagnostic tests. *Biometrics* 36: 167-171 (1980).
- Yanagawa, T., Kasagi, F., and Yoshimura, T. A method to estimate incidence rates of onchocerciasis from skin snip biopsy with consideration of false negatives. *Biometrics* 40: 301-311 (1984).
- Yanagawa, T., and Kasagi, F. Estimating prevalence and incidence of disease from a diagnostic test. In: *Statistical Theory and Data Analysis* (K. Matsushita, Ed.), Elsevier North-Holland Biomedical Press, New York, 1985, pp. 783-799.
- Gastwirth, J. L. Statistical precision of medical screening procedures. *Stat. Sci.* 2(3): 213-222 (1987).
- Yanagawa, T., and Tokudome, S. Methodological aspects in estimating the disease risk of adult T-cell leukemia/lymphoma. *Bull. Biometric Soc. of Jpn.* 9(1): 29-37 (1988).
- Thomas, D. G., and Gart, J. J. A table of exact confidence limits for differences and ratios of two proportions and their odds ratios. *J. Am. Stat. Assoc.* 72: 73-76 (1977).
- Katz, D., Baptista, J., Azen, S. P., and Pike, M. C. Obtaining confidence intervals for the risk ratio in cohort studies. *Biometrics* 34: 469-474 (1978).
- Santner, T. J., and Snell, M. K. Small sample confidence intervals for P_1-P_2 and P_1/P_2 in 2×2 contingency tables. *J. Am. Stat. Assoc.* 75: 386-394 (1980).
- Koopman, P. A. R. Confidence intervals for the two binomial proportions. *Biometrics* 40: 513-517 (1984).
- Bedrick, E. J. A family of confidence intervals for the ratio of two binomial proportions. *Biometrics* 43: 993-998 (1987).
- Gart, J. J., and Num, J. Approximate intervals estimation of the ratio of binomial parameters: a review and corrections for skewness. *Biometrics* 44: 323-338 (1988).
- Gart, J. J. The analysis of ratios and cross-product ratios of Poisson variates with application to incidence rates. *Commun. Stat.* A7: 917-937 (1978).
- Takatsuki, K., Uchiyama, T., Sagawa, K., and Yodoi, J. Adult T-cell leukemia in Japan. In: *Topics in Hematology* (S. Seno, F. Takaku, and S. Irino, Eds.), Excerpta Medica, Amsterdam, 1977, pp. 73-77.
- Tokudome, S., Tokunaga, O., Shimamoto, Y., Miyamoto, Y., Sumida, I., Kikuchi, M., Takeshita, M., Ikeda, T., Fujiwara, K., Yoshihara, M., Maeda, Y., Yanagawa, T., and Nishizumi, M. Incidence of adult T-cell leukemia/lymphoma among human T-lymphotropic virus type I carriers in Saga, Japan. *Cancer Res.* 49: 226-228 (1989).
- Hinuma, Y., Nagata, K., Hanaoka, M., Nakai, M., Matsumoto, T., Kinoshita, K., Shirakawa, S., and Miyoshi, I. Adult T-cell leukemia: antigen in an ATL cell line and detection of antibodies to the antigen in human sera. *Proc. Natl. Acad. Sci. USA* 78: 6476-6480 (1981).
- Ikeda, M., Fujino, R., Matsui, T., Yoshida, T., Komoda, H., and Imai, J. A new agglutination test for serum antibodies to adult T-cell leukemia virus. *Gann* 75: 845-848 (1984).
- Kamihira, S., Nakashima, S., Ichimaru, M., Moriuchi, Y., Oyakawa, Y., Okuda, H., Kanamura, M., and Oota, T. The problems concerning assay of anti-ATLA by immunofluorescence, enzyme-immuno assays, gelatin particle agglutination and Western blot methods. *J. Jpn. Soc. Blood Trans.* 33(1): 7-11 (1987).
- Yoshida, T., and Yamamoto, N. ATL virus antibody (ATLA) koutai no sokuteiho [in Japanese]. *Nippon Rinsho* 44(10): 132-138 (1986).